# 重大突发公共卫生事件驱动的医疗数据开放治理模式研究*

■ 李嘉兴[1] 王雷[2] 宋士杰[3] 朱庆华[2]
[1] 南京农业大学信息管理学院 南京 210095 [2] 南京大学信息管理学院 南京 210023
[3] 河海大学商学院 南京 211100

摘 要：[目的/意义]通过深入探析重大突发公共卫生事件医疗数据开放共享过程，为我国重大突发公共卫生事件数据治理构建一种多主体协同参与的医疗数据开放治理模式。[方法/过程]采用文献调研和网络调查相结合的方法，全面梳理国内外重大公共卫生事件医疗数据开放模式的研究现状和现存障碍，从解决实际问题的角度提出治理方案。[结果/结论]通过明确重大突发公共卫生事件医疗数据开放原则，分解数据开放共享过程，提出对重大突发公共卫生事件下医疗数据开放共享进行分阶段治理的模式，并制定相关配套保障机制，最终提出一套切实可行的面向重大突发公共卫生事件的医疗数据开放治理模式。

关键词：重大突发公共卫生事件 健康大数据 数据开放 数据治理模式
分类号：G251.2
DOI：10.13266/j.issn.0252-3116.2022.04.003



## 1 引言

新型冠状病毒肺炎疫情（以下简称“新冠疫情”）已在全世界范围内持续传播超过两年时间，新冠疫情在蔓延过程中不断产生新的变异毒株，对疫情防控造成巨大压力。习近平总书记指出这次新冠疫情是新中国成立以来传播速度最快、感染范围最广、防控难度最大的一次重大突发公共卫生事件[1]。

新冠疫情等重大突发公共卫生事件具有突发性、公共性及不确定性等特征，防控难度极高且易引发民众的恐慌情绪[2]。因此在应对重大突发公共卫生事件时应保证及时、准确、全面、透明、公开相关医疗数据，从而提升国家形象，增强全社会抗疫信心，有效抑制疫情传播。但是，重大突发公共卫生事件医疗数据开放的高标准使常规医疗数据开放模式面临严峻挑战，现有模式与理想模式之间的巨大差距亟需弥补。

针对重大突发公共卫生事件医疗数据开放这一重大问题，习近平总书记在疫情防控工作会议上指出，防控疫情要加大科研攻关力度，推动疫情相关数据和病例资料的开放共享，及时完善防控策略和措施[3]。为落实习近平总书记对疫情防控的重大部署，本研究通过梳理医疗数据开放模式现状，总结各国医疗数据开放经验，分析我国重大突发公共卫生事件医疗数据开放的现存障碍，构建重大突发公共卫生事件的医疗数据开放治理模式，提升我国重大突发公共卫生事件医疗数据的开放与共享水平。

## 2 重大公共卫生事件医疗数据开放模式现状

### 2.1 医疗数据开放的一般模式

自 2009 年美国发起开放数据运动以来，开放数据运动在全球范围内迅速兴起，根据万维网之父 T. B. Lee 创立的万维网基金会发布的全球开放数据晴雨表最新版排名，英国和加拿大并列第 1 位，美国作为开放数据的发起国位于第 9 位，而中国的排名是 24 位，说明中国的数据开放程度和世界上的数据开放强国仍有差距。医疗作为优先的民生领域，美国、英国、加拿大和中国均将其作为优先开展数据开放的领域，接下来将对这 4 个国家医疗领域的数据开放模式现状进行

* 本文系国家自然科学基金面上项目“社交媒体环境下失真健康信息的传播机制与协同治理研究”（项目编号：72174083）研究成果之一。
作者简介：李嘉兴，讲师，博士；王雷，博士研究生；宋士杰，讲师，博士；朱庆华，教授，博士生导师，通信作者，E-mail：qhzhu@nju.edu.cn。
收稿日期：2021-10-07 修回日期：2021-12-15 本文起止页码：23-32 本文责任编辑：徐健

说明。

美国是最早推出政府数据开放平台(data.gov)的国家,随着专门针对健康医疗的数据开放门户 healthdata.gov 的推出,进一步加速了美国医疗数据的开放,目前美国政府已发布的健康医疗领域开放数据资源数量全球最多。从数据来源机构来看,以 healthdata.gov 为例,主要有医疗保险与医疗补助服务中心、疾病预防控制中心和国家食品药品监督管理局等[4]。与数据发布机构的职能相对应,数据的主题以医疗保险与补助相关数据居多(36%),其次为卫生费用相关数据(11%)[5]。从数据的更新频率来看,这些医疗健康数据大多都是以年为单位进行统计,但有些年份的数据存在缺失。在开放数据平台的特征方面,healthdata.gov 不仅提供数据检索及多维度分类导航,还有开发者接口及问题反馈板块。基于 T. B. Lee 所推荐的五星发展模型对数据开放进行评级,healthdata.gov 平台中的数据有近 80% 为访问受限的数据(0 星),可开放获取的健康医疗数据多为结构化数据(3 星)[5]。

对于政府开放数据准备度常年位于全球最佳的英国来说,在健康医疗领域,政府主导建设了全国性项目国家医疗服务系统,该系统建立了可在信息中嵌入高质量标准经验的线上平台、使公民和社会团体可评估政府工作进展的认证机制[6]。英国健康医疗领域的开放数据门户包括 data.gov.uk 和 NHS Choices 等。在数据的来源机构方面,在 data.gov.uk 中,健康医疗数据集的主要来源机构是健康和社会保健信息中心、卫生部和苏格兰信息服务部等[7],但健康医疗数据的主题是以卫生管理与质量监测类数据(38%)、人口统计数据(32%)为主。此外,大多数医疗开放数据的更新没有固定的频率。在开放数据平台的特征方面,英国的医疗开放数据平台提供数据发布入口、操作手册和 API 接口,便于数据发布者及使用者快捷发布、获取和利用数据。在医疗数据的开放程度方面,有近 70% 为访问受限的数据(0 星),可开放式获取的数据也多为结构化数据[5]。

加拿大的第一个开放数据统一网站 www.open.canada.ca/en/于 2011 年建立,加拿大政府还与 Health Infoway 公司合作建立了覆盖全加拿大的电子健康系统,该电子健康系统向加拿大全国居民提供准确的个人电子医疗记录,医疗机构亦在授权后获得该纪录,患者注册及管理可以实现跨省操作,并且共享信息的隐私及安全能够得到保障[8]。加拿大医疗开放数据的主要来源机构包括加拿大食品检验局、加拿大卫生部和加拿大公共卫生局等。加拿大的医疗开放数据主题主要包括食物营养类、疾病病症类和健康生活类。在开放数据的更新频率方面,加拿大的医疗开放数据同样没有固定的更新频率。在开放数据平台的特征方面,加拿大的开放数据平台注重用户体验,融合社交媒体功能,支持评论、提议自己感兴趣的数据资源、上传 APP 等多种形式的用户互动[9]。

我国作为参与数据开放较晚的国家,也在医疗数据的开放方面进行了一些探索和实践。上海市在 2012 年率先上线了“政务数据服务网”,开启了我国政府数据开放平台的序幕。国务院 2015 年在《促进大数据发展行动纲要》中提出要建成全国政府数据统一开放平台[10],并优先在交通、医疗、卫生等领域实现公共数据资源合理适度开放,目前平台已经在 50 多个地市投入了使用。在医疗健康数据的来源机构方面,以上海市公共数据开放平台为例,主要包括市市场监管局、市卫生健康委、市统计局等。在医疗健康数据的主题方面,我国的数据开放平台中主要包括公共卫生类数据(卫生直报数据共享和社区慢性病健康管理数据共享等)[11]、临床类数据(就诊治疗信息查询、科研数据获取和医联体平台建设等)[12-13],以及医保类数据(医保信息与医院信息的交互等)。而对于开放数据的更新频率,大多数据集是以年为单位,但有些年份的数据存在缺失。在开放数据平台的特征方面,包括数据的分类检索、数据可视化和互动社区等模块,还面向实际应用场景提供了一些行业应用案例。此外,基于这些医疗开放数据国内还开展了很多学术研究,包括理论策略的提出(建立共享数据目录、平台制度建设等)[9,14],以及技术方案的研究(区块链技术、云计算等)[15-16]。

可以看出,美国、英国、加拿大均已建立了全国统一的数据开放平台,医疗开放数据的来源机构和主题则各有不同,而数据的更新大多数没有固定的频率。此外,在这些国家的医疗数据开放网站中,除了检索和查询功能,还加入了很多评论、互动和反馈等模块以增加用户的参与感。而我国由于数据开放起步较晚,全国性的数据统一开放平台仍在建设和推广中,很多地市还没有固定的数据开放平台,但上海、北京这些地区的医疗开放数据资源丰富,遍及医院内部数据和社会上的医保健康等数据,开放数据平台也在传统的检索、查询功能基础上加入了用户互动模块,还提供一些面向应用场景的案例。

## 2.2 重大公共卫生事件的医疗数据开放模式

新冠疫情的出现引起了人们对重大公共卫生事件

中数据开放的关注。M. Laessig 曾提出开放的医疗健康数据在重大公共卫生事件中起着重要作用,从疫情暴发和大流行的早期预警,到为个人提供个性化药物,再到支持卫生系统的管理,疫情防治的每个阶段都离不开海量开放医疗数据的支撑[17]。此外,在早期的埃博拉病毒[18]和寨卡病毒[19]暴发过程中,医疗健康数据的开放已经被证明非常有效。

与新冠疫情相关的医疗开放数据主题主要包括感染病例、政策、人口流动、社交媒体和科研文献 5 个方面。

感染病例方面的数据开放平台包括 2019-nCoV 时间序列感染数据仓库,这是丁香园在新冠疫情暴发后开放的数据集之一,专注于中国报告的病例,其基于 Python 爬虫每小时对数据进行更新。约翰霍普金斯大学系统科学与工程中心也发布了新冠病毒数据存储库[20],其基于 Web 的交互式功能实时追踪全球新冠病例。中国、美国、英国、日本等国家都报告了其二级行政区划的新冠病例[21]。GNU Health 是通过提供流行病学监测报告以研究传染性疾病的发病率和流行率的数字健康生态系统,且已经和欧盟委员会、阿根廷政府进行合作,应用在新冠疫情的治理中[22]。

政策相关的数据开放方面,牛津的政府新冠响应追踪器系统地收集不同国家政府的政策响应信息,并对这些措施的严格程度进行评分,提供了一种跨国、跨时间的政策展示平台,也用于了解政府应对措施在疫情暴发期间是如何演变的。凯撒家庭基金会则发布了美国的州政策行动,包括居家令、旅行者隔离、企业停业、大型聚会禁令、学校停课等。

人口流动数据,即根据一些定位设备以及行程信息跟踪人们的移动获得的数据,这有助于探索病毒传播的空间趋势。弗吉尼亚大学 Dataverse 发布了人口流动数据,使用去识别化和聚合的手机数据、航空旅客行程数据和病例报告以评估新冠病毒在中国的传播风险。笛卡尔实验室的 M. S. Warren 和 S. W. Skillman 使用匿名移动设备位置来衡量病毒的传播风险[23]。尽管新冠肺炎接触者的追踪在传染病调查中很重要,但移动数据的公开不应包含个人可识别数据,世卫组织也发布了接触者追踪应用的道德指南,禁止使用公共卫生部门生成的个人可识别数据[24]。

与新冠疫情相关的社交媒体数据既可以揭示疫情方面的错误信息和未经证实的谣言,也有助于了解公众的恐惧、恐慌和其他社会情绪。其中一个与新冠疫情相关的 Twitter 数据集[25]包含自 2020 年 1 月 22 日以来收集的推文,这些推文基于冠状病毒、covid-19、covid19、sars-cov-2 等关键字,其还跟踪 CoronaVirusInfo、V2019N、WHO 等相关帐户,该数据集涵盖了来自世界各地的推文,但不包含用户的地理位置信息。学者们基于此数据集,可以对新冠疫情相关谣言的传播以及疫情期间人们的情绪等问题展开分析。

为了应对新冠疫情大流行,世卫组织和艾伦人工智能研究所都公开了有关新冠疫情的免费学术文章资源[20],供全球研究界使用,艾伦人工智能研究所提供的资源包括超过 51 000 篇文章,这类开放数据旨在动员研究人员应用自然语言处理进行分析,以产生新的见解支持抗击新冠疫情。

### 2.3 模式对比分析

将重大公共卫生事件下的医疗数据开放模式与一般的医疗数据开放模式相对比,主要有以下几种区别:

#### 2.3.1 重大公共卫生事件下更加强调医疗数据开放的实时性

在重大公共卫生事件中,整个社会面临的形势更严峻,变化也更频繁。因此,提供更具实时性、时效性的数据,才可以更准确地分析当下的疫情,也可以基于时间维度了解疫情、政府政策和民众情绪等方面的演变。在医疗数据开放的一般模式中,即使是美、英、加这些数据开放程度较高的国家也很难做到实时动态的数据更新,但在新冠疫情中,无论是感染病例还是相关政策的数据开放都体现了数据开放的实时性。

#### 2.3.2 重大公共卫生事件下的医疗数据开放共享更注重个人隐私保护

个人隐私安全问题一直是医疗数据开放中不可忽略的风险,在重大公共卫生事件中随着数据量增大和数据潜在价值增大,隐私泄露的风险也越大。因此,需要对个人数据做脱敏、去识别化处理后形成不包含隐私的数据集,实现隐私保护和数据开放之间的平衡。新冠疫情的人口流动开放数据以及社交媒体开放数据中均体现了医疗数据开放对于个人隐私的保护。

#### 2.3.3 重大公共卫生事件下的医疗数据开放范围更大

在医疗数据开放的一般模式中,各个国家基于实际情况拥有不同的数据来源,划分不同的数据主题,并且搭建独立的数据开放平台。但重大公共卫生事件往往发生在一个较大的空间范围内,这次的新冠疫情更是遍及全球,因此各国之间需要加强合作,更大范围的数据互通和数据共享有利于共同抗击疫情。目前,感染病例、相关政策和社交媒体数据方面已经出现了一

些跨国甚至全球的数据开放平台，不同国家和地区基于这些平台进行经验分享、学术交流和疫苗研制合作。

## 3 重大公共卫生事件医疗数据开放障碍分析

党的十九届五中全会明确提出加快数字化发展，推动数据资源开发利用[26]。激活重大突发公共卫生事件医疗数据资源的流通，对有效控制重大突发公共卫生事件具有重要战略意义。医疗数据流通过程中的参与主体主要划分为数据控制者及数据使用者，由数据控制者收集、生成医疗数据后提供给数据使用者[27]。因此，本研究将政府及所属部门、医疗机构 3 类具有收集、生产数据能力的主体定义为重大突发公共卫生事件医疗数据控制者，将企业、科研机构、社会公众 3 类具有数据处理能力和数据使用能力的主体定义为数据使用者。我国重大突发公共卫生事件医疗数据流通过程为数据控制者将收集、生产、处理的数据提交至开放平台，数据使用者从开放平台中获取、处理、使用数据，如图 1 所示。这说明了我国重大公共卫生事件医疗数据开放体系由多主体构成，因此本研究将从多主体视角出发，对我国重大突发公共卫生事件医疗数据开放的现存障碍展开分析。

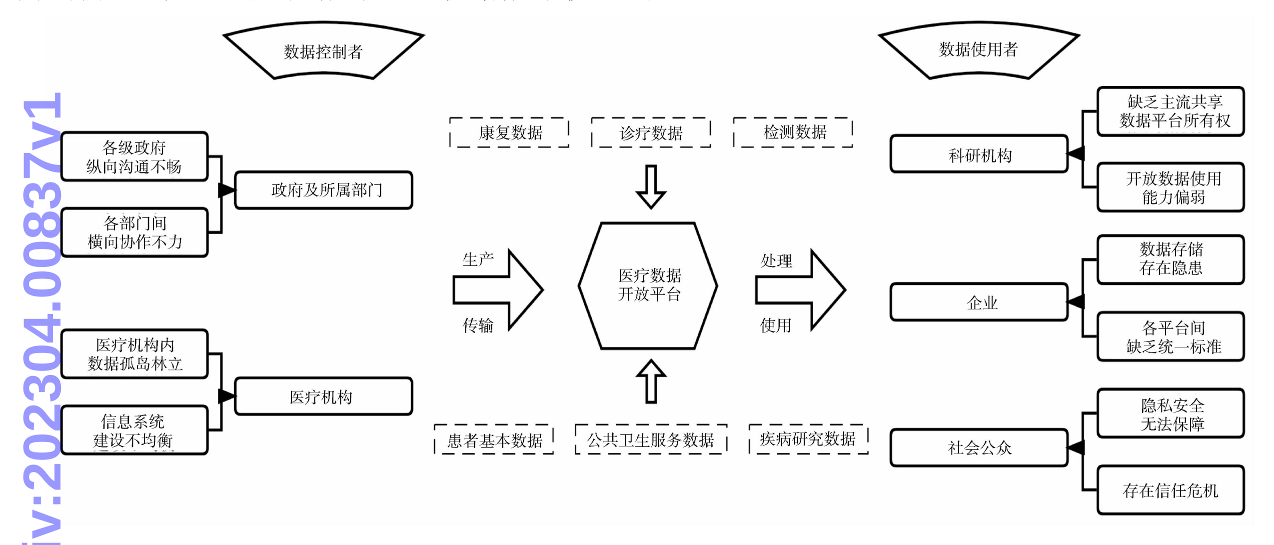



图 1 多主体参与的重大公共卫生事件医疗数据开放体系



### 3.1 数据控制者

#### 3.1.1 政府部门

在我国重大突发公共卫生事件医疗数据开放体系中，政府处于主导地位，部门间协同配合，并调动社会力量广泛参与“大数据抗疫”。在此过程中，政府运用大数据有效支撑服务重大突发公共卫生事件的联防联控，并实现精准施策，体现了大数据在疫情防控中的重要作用。但在重大突发公共卫生事件医疗数据的开放与共享中政府部门也暴露出一些短板。

（1）各级政府纵向贯通不畅。在面临重大突发公共卫生事件时，中央政府担负统筹规划抗疫全局的功能；地方需要落实中央政府的战略，又要担负全域抗疫统筹规划，承担疫情兜底保障工作[28]。在长达两年的抗击疫情过程中，我国纵向治理体系暴露出一些问题，在数据获取方面，地方政府信息化水平存在显著差异，部分地方政府存在瞒报、漏报、错报的情况，导致中央政府无法精准了解疫情状况。在开放数据利用方面，地方政府、基层政府往往只关注本域疫情数据，缺乏全局观念，部分地方政府对有疫情的省市实施“一刀切”原则，对国家经济建设、社会公众日常生活产生一定影响，这就是对开放数据缺乏有效利用，无法精准定位疫情范围的体现。

（2）各部门间横向协作不力。各级政府部门日常处理大量行政事务，在面对重大突发公共卫生事件时便暴露出资源配置不足的短板。各部门往往需要向上级政府、卫健委等同级政府部门传递疫情相关数据，疫情数据的填报工作需要耗费大量人力时间，且重复填报大幅增加了工作量，从而导致部门日常运转出现问题。此外，各政府部门之间数据共享存在壁垒。在面对重大突发公共卫生事件时，政府各部门往往需要等待上级政府整合全面数据，下达疫情防控指令后，各部门再着手准备应对措施，应对措施与疫情发展速度相比产生一定滞后性，对疫情防控产生影响。

#### 3.1.2 医疗机构

医疗机构在重大突发公共卫生事件医疗数据开放体系中占据重要地位，承担着疾病检测、疾病诊断、疾

病治疗的重要任务，能够获取海量一手数据。通过向上级机关报送数据，使中央政府对疫情有清晰、准确地认知，从而提出有效的防控手段。但当前医疗机构间信息化建设水平差异巨大，且医院内部各诊室间呈现松散融合的特征，这种结构特征尚不足以应对重大突发公共卫生事件的冲击。

（1）医疗机构内“数据孤岛”林立。数据孤岛是指数据处于割据状态，从而形成彼此间相互隔绝的数据孤岛[29]。医疗数据不易处理，涉及庞大的不可交互且敏感的文件。以新冠疫情为例，重症患者往往会引发一系列并发症，需要院内各科室的协作诊疗。但是由于医疗数据具有严格的隐私保护政策，患者诊疗数据在同一医院的不同科室往往并不互通，医院内部系统无法实现数据互通共享，易对病患的诊治造成阻碍。

（2）医疗机构间管理信息系统建设不均衡。重大突发公共卫生事件出现后，医疗机构承担疾病诊治及疫情数据报送的任务。我国经济发达地区医疗机构的管理信息系统建设较为成熟，已经进入区域医疗信息化阶段，能够实现区域内医疗机构间、医疗机构与卫生防疫部门、医疗机构内部的信息互联互通，能够有效应对重大突发公共卫生事件，实现精准抗疫、科学防控。而我国经济中等发达地区及经济欠发达地区的医疗机构管理信息系统建设相对落后，区域医疗数据集成平台尚未建成，医疗机构数字化、智能化程度相对较低，应对重大突发公共卫生事件时无法及时实现医疗数据的开放共享，无法满足疫情防控的需要。

### 3.2 数据使用者

#### 3.2.1 企业

当重大突发公共卫生事件来临，我国企业积极承担社会责任，在追踪疫情发展状况、粉碎疫情谣言、引导疫情舆论、推动线上办公等方面作出重要贡献。腾讯、网易、新浪等大型互联网企业通过获取国家、地方政府疫情公开医疗数据，构建疫情实时播报平台，但是各平台间缺乏统一数据标准，疫情医疗数据的全面性、时效性、一致性、可读性无法得到保障，这使得社会公众通过互联网公司开发的疫情数据开放平台获取到的信息不一致，容易对疫情状况产生误判。此外，大量疫情相关医疗数据存放于企业云中，由企业进行数据存储及监管，国家安全及公众隐私安全保障存在一定隐患。

#### 3.2.2 科研机构

科研机构在应对重大突发公共卫生事件中承担病毒基因测序、疫苗研发、疫情预测的重要职能。中国科学院在新冠疫情之初便迅速上线三大科研数据开放共享平台，具体包括 2019 新型冠状病毒信息库、新型冠状病毒肺炎科研文献共享平台、新型冠状病毒国家科技资源服务系统。虽然三大平台收录并发布了大量国内外学者的疫情相关研究成果，但仍不是国际主流的医疗数据共享平台，主流医疗数据共享平台的所有权仍由英美国家牢牢把持[30]。此外，我国科研机构对于疫情的预测能力薄弱，不同研究团队的预测方法差异较大，并没有统一的研究范式对疫情演化进行高精度预测，无法为政府决策提供数据分析支持，这体现了我国科研机构对疫情开放数据的使用能力偏低。

#### 3.2.3 社会公众

在面对重大突发公共卫生事件时，保护社会公众的生命安全是防控的首要工作。在新冠疫情初期，因对疫情认识不全面导致未形成标准化的诊疗过程，使社会公众对地方政府、医疗机构产生信任危机。国家通过大力治理新冠疫情医疗数据开放问题，有效化解此危机，重塑了社会公众对国家疫情防控的信心。而在新冠肺炎诊疗过程中，如何在保障公众知情权的同时，保护患者的隐私安全是另一个亟需解决的问题。因此，面临重大突发公共卫生事件时如何贯彻落实公众隐私安全及提升公众信任感的长效机制还需进一步研究。

## 4 重大突发公共卫生事件驱动的医疗数据开放协同治理模式构建

### 4.1 重大突发公共卫生事件医疗数据开放原则

重大突发公共卫生事件医疗数据开放治理模式构建首先应明确原则，以原则为引导构建数据开放治理模式。2013 年八国集团（G8）发布《数据开放宪章》，对政府数据开放原则作出明确界定，宪章中强调政府数据以开放、改进治理、高质量、数据可为所有人使用和促进创新为原则。在重大突发公共卫生事件背景下，医疗数据开放的原则应进一步厘定。因此，本研究在数据开放宪章的基础上，结合重大突发公共卫生事件背景，提出重大突发公共卫生事件医疗数据开放原则，具体包括责任原则、公开透明原则、质量原则和安全原则。

#### 4.1.1 责任原则

应对重大突发公共卫生事件时，政府掌握几乎全部的医疗数据资源，对医疗数据开放应承担主体责任，医疗及科研机构、企业、公众对医疗数据开放承担次要

责任。重大突发公共卫生事件医疗数据开放需明确各级责任单位在数据开放全流程中的职能定位，明确各级政府及卫生健康行政部门的监管职责[31]。

#### 4.1.2　公开原则

重大突发公共卫生事件医疗数据开放应遵循公开的原则，坚持病毒基因测序数据、病毒溯源数据、脱敏处理后病例数据全面开放，通过医疗数据的开放共享保障公众知情权，维护社会稳定，提升国际形象。

#### 4.1.3　质量原则

真实准确的数据是重大突发公共卫生事件精准决策的坚实基础，是一切抗疫行动的依据。与其他类型开放数据相比，重大突发公共卫生事件医疗开放数据的质量标准更高。因此，重大突发公共卫生事件的医疗开放数据应保证数据完整性、准确性和一致性，基于高质量的医疗数据制定标准化诊疗方案，有效推进特效药研发进展。

#### 4.1.4　安全原则

重大突发公共卫生事件医疗数据涉及海量的公众隐私信息，在数据开放环境下保障患者的隐私信息及公众的健康医疗数据不受侵害对国家安全及社会稳定具有重要意义。在数据开放共享的过程中，应以数据安全法为准绳，贯彻重大突发公共卫生事件医疗数据开放的安全原则。

### 4.2　重大突发公共卫生事件医疗数据开放治理模式总体设计

随着数字技术的发展，数据开放水平不断提升，但在重大突发公共卫生事件背景下全世界对医疗数据开放水平提出更高的要求。当前我国政府、医疗机构、科研机构、企业及社会公众正面临一系列医疗数据开放共享难题。为解决重大突发公共卫生事件医疗数据开放共享障碍因素，本文针对重大突发公共卫生事件医疗数据开放治理模式进行总体设计。

重大突发公共卫生事件医疗数据开放共享过程具有阶段性的特点，具体可划分为数据获取阶段、开放数据阶段和数据共享阶段。数据获取阶段主要包括数据采集、清洗及处理；开放数据阶段涵盖数据公开及管理；数据共享阶段包含数据流动及开发利用。重大突发公共卫生事件医疗数据开放共享是以政府为主导，医疗机构、科研机构、企业协作，公众参与的过程体系，其中涉及多个社会主体且主体之间相互关联。因此，重大突发公共卫生事件医疗数据开放治理应在明确数据开放共享过程的基础上，形成以政府、社会组织及公众为中心的纵向协同治理结构。

重大突发公共卫生事件医疗数据开放治理模式所解决的痛点在于，从数据开放共享过程视角出发对重大突发公共卫生事件的医疗数据获取阶段、开放医疗数据阶段、医疗数据共享阶段进行治理。在数据开放共享的不同阶段对参与主体提出具体的治理策略，有效提升医疗数据隐私性、流通性、时效性、安全性、可利用性，从而缓解我国在重大公共卫生事件背景下的医疗数据开放障碍。

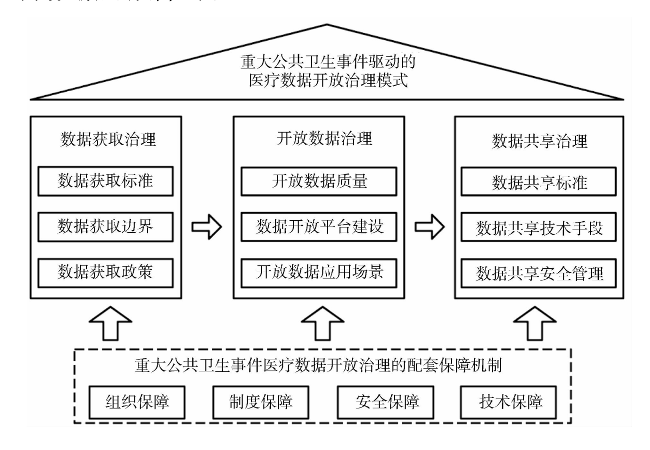



图 2　重大突发公共卫生事件驱动的医疗数据开放治理模式

#### 4.2.1　重大突发公共卫生事件医疗数据获取治理

重大突发公共卫生事件医疗数据开放治理的开端是医疗数据获取治理，目的在于保证突发事件医疗数据的可获取性、完整性和可利用性[32]。数据获取是数据开放共享的基础，在数据获取治理阶段应重点关注数据获取标准及存储安全、加强公众隐私保护等方面。各级政府及所属部门、医疗机构作为医疗数据控制者，应在提升数据获取效率的同时，加强和规范突发事件医疗数据获取过程中产生的一系列安全问题。因此，本研究从 3 个方面对重大突发公共卫生事件医疗数据获取进行治理，分别为制定数据获取标准、明确数据获取边界和完善数据获取政策。

在应对重大突发公共卫生事件时，通过对医疗数据的广泛收集，可以掌握全社会疫情状况和全民健康水平情况。在数据获取过程中，涉及大量公众隐私及社会安全问题，如何标准化获取数据避免"越界"是突发事件医疗数据获取环节亟需解决的问题。突发事件医疗数据采集的标准化是数据存储、加工及数据服务等数据开放治理一系列环节的前提，数据开放治理涉及各级政府及部门、医疗科研机构、企业、社会公众等方面，规范使用突发事件医疗数据资源必须从数据采集的源头来做标准化和规范化。政府作为突发事件医

疗数据获取的引导者,应制定医疗数据获取细则,包括规范获取的数据类型,构建统一的数据获取标准。突发事件医疗数据获取标准仍不清晰,对于各医疗机构、基层政府等收集医疗数据的主要参与者,应制定统一的数据获取标准体系,追求数据的一致性,保障原始数据质量。以数据安全法为基础,加快推动重大突发公共卫生事件医疗数据获取的单独立法,强化各级政府、部门、机构在数据获取中的法律意识,以法律制度为准绳,明确划分突发事件医疗数据采集的合法边界,明晰数据获取过程中参与主体的责权利。加强对数据获取类 APP、小程序的监管,对互联网公司开发的各种医疗数据获取类应用程序必须精细化监管,通过各种形式的审查监测,对擅自超范围获取公众隐私的应用程序加大执法力度,在突发事件医疗数据获取流程上实现可追查化,杜绝一切侵害公众个人隐私的情况发生。

数字技术的发展为重大突发公共卫生事件的防控带来便利,但是并不能以牺牲公众隐私安全为代价获取医疗数据。数据的获取应有边界,通过数据获取阶段治理,对突发事件医疗数据获取进行单独立法,明确实施细则,在边界内合法获取数据,才能够为医疗数据的合理流转打下坚实基础。

#### 4.2.2 重大突发公共卫生事件开放医疗数据治理

在对突发事件医疗数据开放阶段治理中,应重点关注医疗管理信息系统建设不均衡、医疗机构内“数据孤岛”问题,从提升医疗开放数据质量、加强医疗数据开放平台建设、厘清医疗开放数据应用场景等环节着手。突发事件医疗原始数据的质量状况决定了医疗开放数据的质量,提升医疗开放数据质量应从原始数据质量入手,通过设计科学的数据质量评估指标,明确各质量指标权重,建立数据质量监控评估指标体系对原始数据、开放数据进行定期质量评估,实时监控突发事件医疗数据的质量。

在保障突发事件医疗开放数据质量的基础上,加快构建医疗数据管控制度,从数据处理规则、使用规则和数据质量的监测规则等方面入手,对突发事件医疗开放数据进行分级分类处理。将突发事件医疗开放数据以时间为顺序,将数据划分为历史数据、当前数据和未来数据 3 类,通过构建数据处理技术模型对 3 类数据进行处理,不断优化算法以保障医疗开放数据的准确性、一致性和及时性。

开放数据治理的重点环节在于医疗数据开放平台建设及厘清应用场景。医疗数据开放平台应为用户提供基于标准化结构的开放数据服务,通过平台整合医疗开放数据资源,避免医疗数据碎片化,实现跨部门、跨医疗机构的协同,为应对突发事件研发特效药、优化临床诊疗方案提供整体化数据支撑,使医疗数据发挥更大的价值,缓解我国应对重大突发公共卫生事件医疗资源不足的短板。而厘清医疗开放数据应用场景则是为医疗数据共享环节打下基础,面向政府,医疗开放数据应为政府提供决策支持;面向医疗机构,医疗开放数据应为其提供医疗机构间的信息互联互通服务;面向金融保险企业,医疗开放数据应为其提供医保核保应用;面向社会公众,医疗开放数据应提供疫情科普服务、用药查询服务等。

#### 4.2.3 重大突发公共卫生事件医疗数据共享治理

数据共享是数据流通过程的最终环节,是开放数据服务的实施阶段。在重大突发公共卫生事件的背景下,医疗开放数据共享已经成为疫情防控、发展趋势预测、新药研发和平息谣言的重要手段。对于突发事件医疗开放数据共享的治理应围绕制定统一数据共享标准以优化数据流动过程,丰富数据共享技术手段以提升医疗及科研机构数据利用能力,加强数据共享安全管理以增加公众信任感等方面展开。

统一的数据共享标准是支撑重大突发公共卫生事件医疗数据共享的基础。重大突发公共卫生事件医疗数据呈现来源渠道广、共享环节技术标准不统一的特点。因此,上层政府应构建统一的重大突发公共卫生事件医疗数据开放共享通道,制定统一的数据接口及技术标准和实施规范,从而通过整合不同层级政府及部门间的医疗数据资源,贯通政府与社会间的数据共享交换体系。

数据共享技术手段是重大突发公共卫生事件医疗开放数据共享的实现方式,因此治理医疗开放数据共享应大力加强突发事件医疗开放数据共享技术攻关。运用 AI 技术计算机视觉和深度学习技术,对医学影像进行学习训练,辅助医生进行高效、精准的诊断与筛查。通过精细化仿真和预测,缩短特效药开发时间周期,节省新药开发成本。通过知识图谱处理融合病历数据,构建数据关系,促进标准化诊疗方案的提出。使用区块链技术保障数据共享过程全溯源,强化数据共享安全,提升社会公众对开放数据的信任。

数据共享安全管理是重大突发公共卫生事件医疗开放数据共享的保障,因此应对现有的数据安全保护法进一步补充。通过细化数据安全保护条例,明确重大公共卫生事件下医疗开放数据共享过程中的保护对象、范畴和违法责任等。以新冠疫情为例,重大突发公

共卫生事件在全球范围内蔓延,需要全人类共同解决这一难题,因此医疗开放数据共享往往涉及跨境流动的情况,因此制定关于数据开放共享和跨境流动监管的法律条款是必要的。同时拓宽现有数据保护法的范围,将区块链、云计算等新技术新应用场景下的突发事件医疗开放数据保护纳入法律保护范畴。

## 4.3 重大突发公共卫生事件医疗数据开放治理配套保障机制

### 4.3.1 组织保障

面向重大突发公共卫生事件医疗数据开放治理的组织保障具体可划分为两个部分:一是强化各级政府及部门之间的合作;二是促进政府与中国特色高水平智库之间的合作。通过促进跨政府部门合作及政府与高水平智库合作来推动突发事件医疗数据开放治理工作。

目前我国已形成疫情防控跨部门协同机制,但是围绕重大突发公共卫生事件医疗数据开放治理尚未设置特定的跨部门机构。强化各级政府及部门之间的合作可以参考澳大利亚政府数据开放治理的经验[33]。以政府为领导,协调各级卫健委、疾控中心、传染病防控重点实验室等机构共同成立突发事件医疗数据指导委员会,对重大突发公共卫生事件下的医疗数据获取、开放、共享等环节提供专业指导。该机构职能包括监管突发事件医疗数据的存储、利用、共享,上可为中央政府疫情防控决策提供有力支持;下可作为官方信息发布渠道,凭借强大的专业背景平息疫情相关谣言,公开脱敏后的诊疗数据,为公众防疫提供指导并提升公众抗疫信心。

智库在我国治理体系中的作用愈发重要,在重大突发公共卫生事件下智库应为政府防控决策提供"高见、先见、远见、创见"[34]。因此,当前我国亟需围绕重大突发公共卫生事件、重大传染病建设一批中国特色高水平智库。重大公共卫生及重大传染病防控智库建设应吸纳医学科学家、法律专家、数据科学家等专业人士,成立紧密融合的决策支持机构,定期研讨疫情防控进展和阻碍,形成可供政府决策参考的调研报告。

### 4.3.2 制度保障

随着总体国家安全观的提出,数据安全的重要性进一步凸显,我国《数据安全法》于 2021 年 9 月 1 日正式施行,健全了数据安全治理体系。基于此,本文在数据安全法的基础上,结合重大突发公共卫生事件的背景,建立健全重大突发公共卫生事件医疗数据开放治理制度保障,使突发事件医疗数据开放治理有章可循,有法可依。

目前我国关于重大突发公共卫生事件医疗数据安全的立法尚不明确,对突发事件医疗数据安全的分级分类需进一步探索。中央政府应对数据安全法中医疗数据安全相关条款进行补充,强化重大突发公共卫生事件背景下的医疗数据安全保障。各地方政府应在中央政府法律法规的基础上制定适合本地区情况的重大突发公共卫生事件医疗数据开放办法,从而确保突发事件医疗数据安全,维护广大人民群众隐私权,完善国家数据安全保障体系建设。

### 4.3.3 技术保障

当前重大突发公共卫生事件的医疗数据开放正处于探索阶段,患者隐私数据泄露、疫情虚假信息、数据开放滞后等问题亟需解决。为强化我国突发事件医疗数据开放治理,需采用区块链、AI、知识图谱等多种技术手段相结合的方式推进突发事件医疗数据开放共享,具体技术架构如图 3 所示:

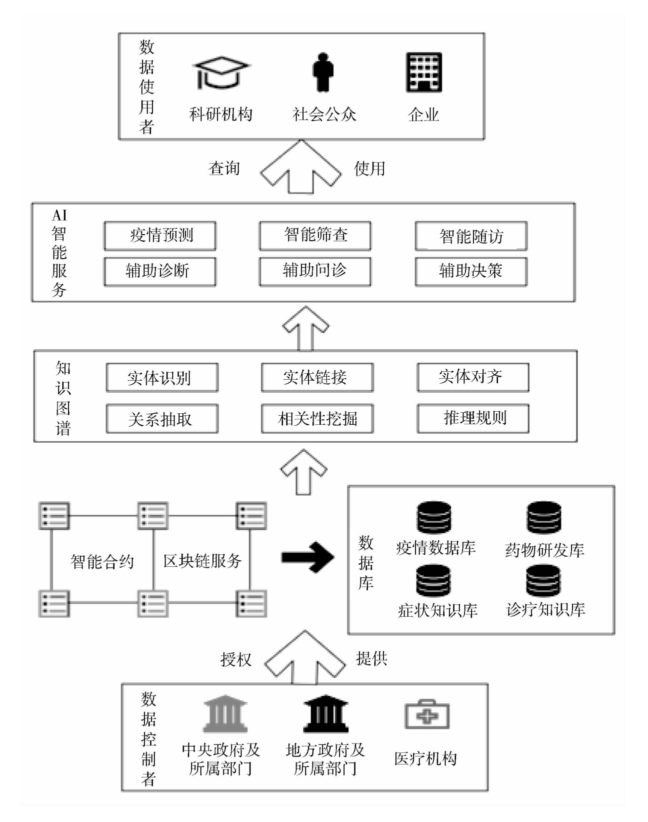



图 3 重大突发公共卫生事件驱动的医疗数据开放技术架构

通过区块链技术搭建重大突发公共卫生事件相关医疗数据库,使数据透明度进一步提升,高透明信息能够有效遏制谣言扩散速度与范围,同时能够有效减少数据错误的产生几率。此外,区块链是建立在信任和

共识基础上的技术，通过突发事件医疗数据共享，配合知识图谱及AI技术提供智能服务，不仅能够有效缓解面临突发事件时医疗资源短缺的短板，还可以有效提升决策速度及准确度。

知识图谱为此次疫情防控提供了有力支持，体现了该技术在重大突发公共卫生事件医疗数据开放共享的重要价值。知识图谱对多源异构数据进行深度提炼，使原本繁杂的医疗数据间形成关联，运用知识图谱对突发事件医疗数据进行处理，能够快速识别数据间的关联，对重大突发公共卫生事件防控、预测、疫苗研发、疾病溯源等环节提供决策支持。而且通过与区块链技术的结合，可以有效解决数据可信度的问题，能够进一步明确在治理过程中各参与主体的权责。

通过上文对区块链、AI、知识图谱等技术手段的研究可以发现，三者间具有极强的互补性，配合度极高。因此，对突发事件医疗数据开放进行治理的过程中，新技术手段应配合使用，才能实现治理效能最大化。

## 5 结语

重大突发公共卫生事件医疗数据开放治理应坚持以政府为主导，各部门、医疗科研机构、企业和社会公众协同配合。在治理过程中遵循责任原则、公开原则、安全原则、质量原则。在突发事件医疗数据开放过程中，以保护人民群众权益为中心，不断优化突发事件医疗数据开放过程，细化数据开放治理模式，加强数据处理新技术在数据开放治理中的应用。通过不断完备突发事件医疗数据开放治理模式，将突发事件医疗数据开放治理的成功经验复制到政府开放数据治理及企业开放数据治理等领域，提升我国数据开放治理水平，在实践中逐步探索出一套符合我国国情的数据开放治理模式。



**参考文献：**

[1] 习近平.为打赢疫情防控阻击战提供强大科技支撑[J].实践(思想理论版),2020,28(4):4－6.
[2] 王怡凡,周典,姚辰欢,等.面向重大突发公共卫生事件的医院应急管理弹性能力研究[J].中国医院管理,2021,41(6):12－15.
[3] 人民日报.划重点！防控疫情，习近平要求这么干[EB/OL].[2021－12－12].https://wap.peopleapp.com/article/5105453/5001744.2020-02-04.
[4] MARC D T, ZHANG R, BEATTIE J, et al. Indexing publicly available health data with medical subject headings (MeSH): an evaluation of term coverage[M]//MEDINFO 2015: ehealth-enabled health. Washington:IOS Press,2015:529－533.
[5] 李姣,郭海红,郭珉江.美英政府开放健康医疗数据的主题分布与开放程度量化研究[J].图书情报工作,2015,59(20):132－137.
[6] 刘宁,陈敏.医疗数据开放方法及策略研究[J].中国医院管理,2015,35(9):37－39.
[7] BRIGHT J, MARGETTS H Z, NING W, et al. Explaining usage patterns in open government data: the case of data.gov.UK[J]. SSRN electronic journal,2015,24(1):1－22.
[8] 杨佳泓,张建国,母晓莉.国际区域医疗信息共享系统建设的现况研究[J].中国医院,2010,14(10):4－6.
[9] 李赞梅.中、英、美、加、澳五国健康医疗领域政府开放数据的现状与启示[J].中华医学图书情报杂志,2019,28(6):42－47.
[10] 国务院印发《促进大数据发展行动纲要》[J].中国卫生信息管理杂志,2015,12(5):447.
[11] 毛云鹏.卫生统计直报数据共享的成功案例[C]//中国卫生信息学会.中国卫生信息技术交流大会暨两岸四地卫生信息化交流会议论文集，大连:中国卫生信息管理杂志社,2012:304－307.
[12] 韦力,梁悦,刘宇宏.基于医疗数据发布中心的患者病历获取服务[J].中国医院管理,2018,38(11):40－41.
[13] 史华新,刘保延,谢琪.中医医疗与临床科研信息共享系统数据采集平台服务流构建的思考[J].中国中医药信息杂志,2013,20(8):4－5.
[14] KASS-HOUT T A, GRAY S K, MASSOUDI B L, et al. NHIN, RHIOs, and public health[J]. Journal of public health management and practice,2007,13(1):31－34.
[15] 黄娜娜,沈亮,杨元原.医疗云中高效安全的数据共享方案研究[J].计算机工程与应用,2020,56(1):92－97.
[16] 周辉,王丽丹,钟成跃.区块链助力电子医疗数据共享[J].解放军医院管理杂志,2019,26(7):645－647.
[17] LAESSIG M, JACOB B, ABOUZAHR C. Opening data for global health[M]. The palgrave handbook of global health data methods for policy and practice. London: Palgrave Macmillan, 2019.
[18] YOZWIAK N L, SCHAFFNER S F, SABETI P C. Data sharing: make outbreak research open access[J]. Nature news,2015,518(7540):477－479.
[19] KIENY M-P, MOORTHY V, BAGOZZI D. Use open data to curb zika virus[J]. Nature,2016,533(7604):469.
[20] DONG E, DU H, GARDNER L. An interactive web-based dashboard to track COVID-19 in real time[J]. The lancet infectious diseases, 2020, 20(5):533－534.
[21] HU T, GUAN W W, ZHU X, et al. Building an open resources repository for COVID-19 research[J]. SSRN electronic journal, 2020,4(3):130－147.
[22] KOBAYASHI S, FALCÓN L, FRASER H, et al. Using open source, open data, and civic technology to address the COVID-19 pandemic and infodemic[J]. Yearbook of medical informatics, 2021,30(1):38－43.
[23] WARREN M S, SKILLMAN S W. Mobility changes in response to

COVID-19[EB/OL]. [2021-12-24]. https://arxiv.org/pdf/2003.14228.pdf.

[24] ORGANIZATION W H. Ethical considerations to guide the use of digital proximity tracking technologies for COVID-19 contact tracing: interim guidance, 28 May 2020[R]. World Health Organization, 2020.

[25] BANDA J M, TEKUMALLA R, WANG G, et al. A large-scale COVID-19 twitter chatter dataset for open scientific research-an international collaboration[J]. Epidemiologia, 2021, 2(3): 315-324.

[26] 新华网. 中国共产党第十九届中央委员会第五次全体会议公报[EB/OL]. [2021-09-14]. http://www.xinhuanet.com/politics/2020-10/29/c-1126674147.htm. 2020-10-29(1).

[27] 高富平. 数据流通理论 数据资源权利配置的基础[J]. 中外法学,2019,31(6):1405-1424.

[28] 赵娟,孟天广. 数字政府的纵向治理逻辑:分层体系与协同治理[J]. 学海,2021(2):90-99.

[29] 周茂君,潘宁. 赋权与重构:区块链技术对数据孤岛的破解[J]. 新闻与传播评论,2018,71(5):58-67.

[30] 崔宇红,王飒. 新型冠状病毒突发公共卫生事件中的数据共享机制研究[J]. 图书情报工作,2020,64(15):104-111.

[31] 国家卫生健康委员会.《关于印发国家健康医疗大数据标准、安全和服务管理办法(试行)的通知》[EB/OL]. [2021-11-24]. http://www.cac.gov.cn/2018-09/15/c_1123432498.htm.

[32] 顾立平. 数据治理——图书馆事业的发展机遇[J]. 中国图书馆学报,2016,42(5):40-56.

[33] 陈美. 澳大利亚地方政府开放数据的保障机制研究——基于多元公共行政观的视角[J]. 情报理论与实践,2017,40(12):139-144,111.

[34] 杨亚琴. 中国特色新型智库现代化建设的若干思考——以智库影响力评价为视角的分析[J]. 中国科学院院刊,2021,36(1):70-78.

**作者贡献说明:**
李嘉兴:论文撰写、修改及论文最后版本修订;
王雷:收集文献、翻译摘要及论文校对;
宋士杰:论文修改;
朱庆华:论文框架设计。



## Research on Open Governance Model for Healthcare Data Driven by Major Public Health Emergencies

Li Jiaxing[1] Wang Lei[2] Song Shijie[3] Zhu Qinghua[2]
[1] School of Information Management, Nanjing Agricultural University, Nanjing 210095
[2] School of Information Management, Nanjing University, Nanjing 210023
[3] Business School, Hohai University, Nanjing 211100

Abstract: [**Purpose/significance**] Through an in-depth analysis of the process of the healthcare data opening and sharing in major public health emergencies, this paper proposes an open governance model of healthcare data with the collaboration participation of multiple subjects for data governance of major public health emergencies in China. [**Method/process**] By using a combination of literature research and Internet surveys, this study comprehensively reviewed the current research status and existing obstacles of healthcare data opening models for major public health emergencies at home and abroad, and proposed governance solutions from the perspective of solving practical problems. [**Result/conclusion**] By clarifying the principles of healthcare data opening for major public health emergencies, this study decomposes the process of data opening and sharing, conducts staged governance model of healthcare data opening and sharing under major public health emergencies, and formulates relevant supporting security mechanisms. Finally, this study proposes a practical governance model of healthcare data opening in major public health emergencies.

Keywords: major public health emergencies  healthcare big data  data opening  data governance model